

贝叶斯理论在遗传概率计算中的应用

河北省保定市第一中学(071000) 龚军辉

摘要 在解决遗传概率问题时,运用贝叶斯理论,能够更加精确地估算发病风险。讨论了贝叶斯理论在多种情况下的应用,为教学作有价值的推广。

关键词 遗传概率;计算;贝叶斯理论

文章编号 1005-2259(2017)10-0041-03

1 问题的提出

中国中小学教育教学网(www.k12.com.cn)论坛上有一个关于遗传概率计算题的讨论,在广大高中生物教师中引起了很大的争议。题目如下。

例1 图1是某家族单基因遗传病系谱图(相关基因用A,a表示),已知II-5是纯合体。请据图回答:若III-12与III-13个体婚配,则子女患病的概率是多少?

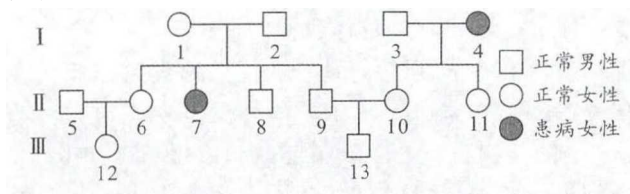


图1

该题争论的焦点在于II-9个体的基因型是AA或Aa的概率分别是多少。大多数教师认为II-9个体的基因型及概率为 $1/3AA$ 和 $2/3Aa$ 。不少文献对本题作了讨论,并尝试利用贝叶斯(Bayes)理论对该题的概率问题进行了分析和解答^[1-4]。

3 关于遗传题命题的思考与备考建议

应该说,江苏遗传题的命制非常鲜明地突出了“新材料、新情境、新思维”的特点,重点考查了考生的信息提取与加工、材料分析、思维创新等能力,试题水平得到进一步的提升,对考生和教师也提出了更高的要求。

鉴于此,在今后的备考环节中,我们应当努力做到以下三点要求。第一,要有积极稳健的心理素质。面对此类题目,考生应当处变不惊,不能被一些陌生的新名词、新环境所困扰,努力适应试题情境。第二,加强思维能力的培养。从本题可以看出,高考对考生

在解决遗传概率问题时,运用贝叶斯理论,能够更加精确地估算发病风险,因此贝叶斯理论在遗传概率的计算中有着特别重要的作用。但由于长期以来,很多教师对遗传概率的计算普遍有一个根深蒂固的认识误区,从而导致如上题所述的问题时至今日还没有统一的结果,争论还在持续。因此,笔者认为有必要对这一类概率问题和贝叶斯理论再进行剖析。

2 贝叶斯理论

贝叶斯理论是一种确认两种相互排斥事件相对概率(如在常染色体隐性遗传病的条件下,正常个体是显性纯合体或者是携带者的概率)的理论,它是在知道了结果的情况下来反推原因事件的概率,是在事后对原因发生可能性大小的重新认识。因此,在遗传概率计算时,根据各项条件推算携带者的概率,可更为准确地计算出各遗传病的发病概率^[5]。

贝叶斯理论涉及以下几个概念。

前概率 $P(B_i)$,指按照遗传规律或系谱特点得出的理论概率,即根据孟德尔遗传规律得出的理论概率。

思维逻辑要求非常高,很多学生因为无法理解所给定的实验步骤而无法入手,因此,在平时训练中一定要加强对思维能力的培养。第三,努力追求“1分”。在实际的考试中,无论最后的压轴题的难度如何,开始的几个小设问一般难度不大。因此,在解题当中要努力追求每一分,最终开辟出一个属于自己的理想天地。

参考文献

- [1] 赵素然.果蝇的保种技巧[J].实验室科学,2005(4):69.
- [2] 戴灼华,王亚馥,栗翼玫.遗传学[M].2版.北京:高等教育出版社,2008.
- [3] 赵寿元.现代遗传学[M].北京:高等教育出版社,2001. △



条件概率 $P(A|B_i)$, 指在事件 A 已经发生的前提下, 某一特定原因出现的概率。条件概率需要从系谱中提供的遗传信息来确定, 如已知家庭成员的健康状况, 正常子女数等。

联合概率 $P(B_i)P(A|B_i)$, 指某一种基因型前提下前概率和条件概率所说明的两个事件同时出现的概率, 即前概率和条件概率之积。

后概率 $P(B_i|A)$, 指每一种基因型下的联合概率除以各事件联合概率之和。即:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum P(B_i)P(A|B_i)}$$

在做遗传概率计算时, 很多教师的认识误区就在于没有区分前概率和后概率。

3 贝叶斯理论的具体应用

贝叶斯理论在遗传概率计算中的应用, 主要是在双亲之一或双方的基因型未知的情况下, 即可应用贝叶斯理论分别计算出所要知道的后概率, 从而计算出正常概率或发病风险。用贝叶斯理论所计算出的概率, 比起只用一般的遗传规律得出的概率更准确, 更接近实际。

3.1 例 1 的分析

现用贝叶斯理论来解决例 1 中 III - 12 与 III - 13 个体婚配后子女患病的概率, 其中最关键的是要确定 III - 12 和 III - 13 个体基因型为 Aa 的概率。

由题意可知, II - 9 个体的基因型为 AA 或 Aa, II - 10 个体的基因型为 Aa。

前概率: 理论上 II - 9 个体基因型为 AA 的概率是 1/3, 为 Aa 的概率是 2/3 (此即为大多数教师的观点)。

条件概率: 如果 II - 9 个体基因型为 AA, 则 II - 9 和 II - 10 生出正常个体的概率为 1; 如果 II - 9 个体的基因型为 Aa, 则 II - 9 和 II - 10 生出正常个体的概率为 3/4。

联合概率: II - 9 个体基因型为 AA 情况下的联合概率为 $1/3 \times 1 = 1/3$, II - 9 个体基因型为 Aa 情况下的联合概率为 $2/3 \times 3/4 = 1/2$ 。

后概率: II - 9 个体的基因型为 AA 的概率是 2/5, 为 Aa 的概率是 3/5。分析见表 1。

表 1 II - 9 个体的概率分析

	II - 9 是 AA	II - 9 是 Aa
前概率	1/3	2/3
条件概率	1	3/4
联合概率	$1/3 \times 1 = 1/3$	$2/3 \times 3/4 = 1/2$
后概率	$\frac{1/3}{1/3 + 1/2} = 2/5$	$\frac{1/2}{1/3 + 1/2} = 3/5$

因此, II - 9 个体为 AA 的概率是 2/5, 为 Aa 的概率是 3/5。II - 10 个体基因型为 Aa, 则 II - 9 和 II - 10 婚配后所得 III - 13 的两种情况是: AA 基因型的概率为 $2/5 \times 1/2 + 3/5 \times 1/3 = 2/5$, Aa 基因型的概率为 $2/5 \times 1/2 + 3/5 \times 2/3 = 3/5$ 。

用同样的方法可以确定 II - 6 各基因型的概率, 分析见表 2。

表 2 II - 6 个体的概率分析

	II - 6 是 AA	II - 6 是 Aa
前概率	1/3	2/3
条件概率	1	1
联合概率	$1/3 \times 1 = 1/3$	$2/3 \times 1 = 2/3$
后概率	$\frac{1/3}{1/3 + 2/3} = 1/3$	$\frac{2/3}{1/3 + 2/3} = 2/3$

因此, II - 6 个体为 AA 的概率为 1/3, 为 Aa 的概率为 2/3。II - 5 个体基因型为 AA, II - 5 和 II - 6 婚配后所得 III - 12 个体基因型为 AA 的概率是 $1/3 + 2/3 \times 1/2 = 2/3$, 为 Aa 的概率是 $2/3 \times 1/2 = 1/3$ 。

因此, III - 12 和 III - 13 婚配, 生出患病后代的概率为 $3/5 \times 1/3 \times 1/4 = 1/20$ 。

3.2 伴性遗传的实例

例 1 是考虑常染色体遗传的例子, 下面再举一个伴性遗传的例子, 以更好地说明贝叶斯理论在解决遗传发病率中的准确性^[6]。

例 2 图 2 为一个色盲家系的遗传图谱 (相关基因用 A, a 表示)。计算 III - 13 和 III - 14 婚配后, 生出患病后代的概率是多少?

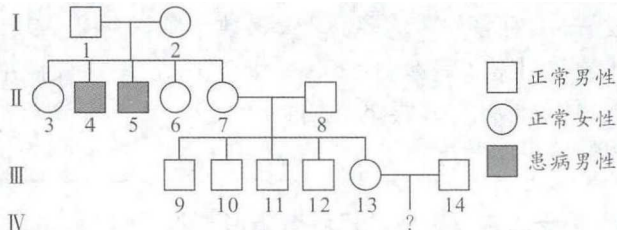


图 2

先按照常规方法来推算: 易知 I - 1 和 I - 2 的基因型分别为 $X^A Y$ 和 $X^A X^a$, 则 II - 7 个体的基因型是 $X^A X^a$ 的概率为 1/2, 从而 III - 13 是 $X^A X^a$ 的概率为 1/4, 因此, III - 13 和 III - 14 婚配后, 后代的发病率为 $1/4 \times 1/4 = 1/16$ 。

但是如果按照贝叶斯理论来计算, 结果将有很大差异。由于 II - 7 个体已经生出了 4 个正常的儿子, 所以她是 $X^A X^a$ 的概率已大大降低。下面列表作出推算 (表 3)。

表3 II-7个体的概率分析

	II-7是 $X^A X^A$	II-7是 $X^A X^a$
前概率	1/2	1/2
条件概率	1	$(1/2)^4 = 1/16$
联合概率	$1/2 \times 1 = 1/2$	$1/2 \times 1/16 = 1/32$
后概率	$\frac{1/2}{1/2 + 1/32} = 16/17$	$\frac{1/32}{1/2 + 1/32} = 1/17$

通过推算可以发现,II-7个体是 $X^A X^a$ 的概率已由1/2降低为1/17,III-13个体是 $X^A X^a$ 的概率为1/34。所以,III-13和III-14婚配后生出患病后代的概率为 $1/34 \times 1/4 = 1/136$,已远远低于上述的1/16。

3.3 双亲基因型未知的情况

上述两例均是双亲之一的基因型未知。而在很多遗传家系中会出现双亲的基因型均未知的情况。用贝叶斯理论同样可以解决此类问题。

例3 图3为一个白化病家系的遗传系谱(相关基因用A,a表示)。II-6与II-7已经生出了一个正常的后代。试估算他们生一个患病小孩的风险有多大?

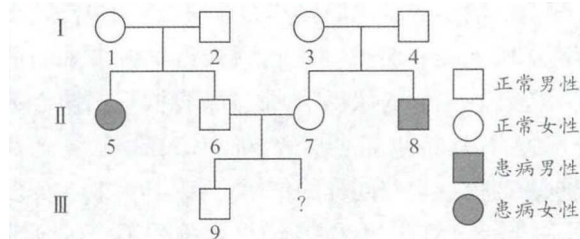


图3

II-6与II-7表现型均正常,并且第一胎已经生出一个正常的儿子,所以无法确定他们的基因型是纯合体AA还是携带者Aa。要求他们生出一个患病孩子的风险,必须估算出他们均为Aa的概率。据图分析,II-6为Aa的前概率为2/3,同理II-7为Aa的前概率也为2/3,所以II-6和II-7均为Aa的前概率为 $2/3 \times 2/3 = 4/9$ 。当他们均为携带者时,生出正常孩子的概率为3/4,因此其条件概率为3/4。另一种情况即II-6与II-7的基因型均为AA或一方为AA,另一方为Aa的前概率为 $1 - 4/9 = 5/9$ 。在这种情况下,生出的孩子均正常,所以其条件概率为1。然后分别计算联合概率和后概率(表4)。

因此,II-6和II-7均为Aa的概率为3/8,则他们生二胎孩子患病的概率为 $3/8 \times 1/4 = 3/32$ 。

表4 II-6和II-7的概率分析

	II-6和II-7均为Aa	II-6和II-7均为AA或一方为AA,另一方为Aa
前概率	$2/3 \times 2/3 = 4/9$	$1 - 4/9 = 5/9$
条件概率	3/4	1
联合概率	$4/9 \times 3/4 = 1/3$	$5/9 \times 1 = 5/9$
后概率	$\frac{1/3}{1/3 + 5/9} = 3/8$	$\frac{5/9}{1/3 + 5/9} = 5/8$

4 推广应用

图4是一个家系的遗传图谱(相关基因用A,a表示)。已知II-4和II-5已经生出了n个正常的后代。试确定II-4基因型是杂合的概率是多少?

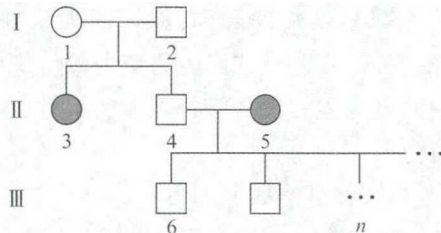


图4

由图易知此病是常染色体隐性遗传病,II-5个体的基因型为aa,II-4个体的基因型可能是AA或Aa,分析见表5。

表5 II-4个体的概率分析

	II-4是AA	II-4是Aa
前概率	1/3	2/3
条件概率	1	$(1/2)^n = 1/2^n$
联合概率	1/3	$2/3 \times 1/2^n$
后概率	$\frac{1/3}{1/3 + 2/3 \times 1/2^n} = \frac{2^{n-1}}{2^{n-1} + 1}$	$\frac{2/3 \times 1/2^n}{1/3 + 2/3 \times 1/2^n} = \frac{1}{2^{n-1} + 1}$

可见,若II-4和II-5已经生出了n个正常的后代,则II-4基因型是Aa的概率为 $1/(2^{n-1} + 1)$,所以当n很大时,Aa的概率就接近于0,而AA的概率就接近于1。

参考文献

- [1] 夏献平.一道概率计算题的分析和思考[J].中学生物教学,2005(6):37-38.
- [2] 王彬,曹岩,欧元亮,等.遗传概率计算题再探讨[J].中学生物教学,2005(1/2):91-92.
- [3] 陈宝明.遗传概率的一个认识误区[J].生物学通报,2007,42(2):36-37.
- [4] 何风华.系谱分析中概率问题的贝叶斯分析一例[J].生物学通报,2007,42(10):25.
- [5] 赵汝良.医学遗传学基础[M].北京:人民卫生出版社,2000:40.
- [6] 陈基平.Bayes理论在遗传咨询中的应用[J].医学理论与实践,2005,18(9):1005-1006.